

10/583940

AP3 Rec'd PCT/PTO 22 JUN 2008

ADVANCED METHOD OF SEARCHING, DRAFTING AND
EDITING OF ELECTRONIC FILES

5 The present invention relates to an advanced method of searching, drafting and editing of electronic files.

10 More particularly, the method according to the invention allows to search electronic documents in specialised data base, within own files and/or in Internet, such as music, texts and videos, according to keys and taking into consideration the searches already carried out, as well as 15 editing said documents in such a way to obtain final documents **for example in a hypertext format**, being sure to have paid the **fees owing for the exclusive rights** for all the parts extracted from the published documents.

15 In the publishing branch of industry, distribution has a large incidence on the **end user price of the work on trade**.

Furthermore, some times it is difficult a precise check of the payment of the copyright fees in case someone partially takes parts of different works. Such a control can be made at present only on works already transferred on a support.

20 However, in this field, we have today the possibility of a wide use of the computer and of the Internet.

However, solutions suggested up to now do not globally improve the situation, rather opening new, restricted market sectors.

25 The **need** thus remains of method and tools able to exploit the informative resources in order to have a global managing of publishing, **thus reducing** manufacturing costs and to offer to the authors the possibility of a quick introduction on the market and to the public a bigger offer with reduced costs.

30 Furthermore, public is interested in the possibility of making advanced searches of documents or of part of documents. However, said searches does not take into consideration the previous searches, i.e. leave in the search engines at most a statistic track and must be set again by each user, starting from the beginning in some cases after the **connect-session's lapse** or at most after 24 hours.

35 Document US 6,078,924 discloses an information platform, which automates the collection of data, provides a method for organising the library of information and provides analysis using multiple-content types, the information platform being a client/server

implementation that is subdivided into four major section: Data Retrieval, Data Classification and Storage, Information Browsing and Desktop Integration. However, method uses a catalogue which can be accessed only by connecting to a central server and searches can be made by roaming the catalogue through different topics and subtopics or by individual information “nuggets”. The adding of a new information source is made by the user through an analysis template guide guiding the user through the data collection, interpretation and analysis process for a specific topics.

10 Thus the need remains of method and tools able to exploit the informative resources in order to carry out advanced searches by advanced search engines. These methods and instruments must be understood as making integral part of the industrial publishing and distribution process and thus have a full industrial character.

15 It is object of the present invention a method for searching, drafting and editing hypertexts, solving the above problems.

It is further specific object of the present invention that of providing apparatuses and instruments necessary for carrying out the method according to the invention.

20 It is **object subject-matter** of the present invention a method of searching, drafting and editing of electronic files comprising the use of one or more peripheral computers, each computer handling an assembly of one or more updatable data bases comprising electronic documents, said electronic documents comprising information suitable to identify the same documents, each peripheral computer handling searches for, drafting of, communication about, and editing of, documents, the method being characterised in that said data base assembly provides one or more catalogues relevant to the documents, the search of documents or of part of documents being carried out employing at least one search criterium or

25 a set of search arguments comprising the use of one or more catalogues

30

Preferably, according to the invention, said one or more catalogues comprise a list of title of the documents

35 Preferably, according to the invention, said one or more catalogues comprise a list of the contexts for which the documents are available, including the **descriptions** of the contexts

Preferably, according to the invention, said one or more catalogues comprise the catalogue of the searches already carried out.

Preferably, according to the invention, said one or more databases provide pointers to a document collection.

5 Preferably, according to the invention, two search criterium are provided, a first search criterium being used in a first step AA, a second criterium being used in a second step BB to semantically analyse documents obtained from step AA.

10 Advantageously, according to the invention, specialised dictionaries are employed in step BB, relevant to specific contexts and/or of the reference semantic assemblies relevant to the contexts.

Advantageously, according to the invention, said assembly of one or more databases is identical for all the peripheral computers.

15 Advantageously, according to the invention, said information suitable to identify documents is text information.

Preferably, according to the invention, said documents are hypertext documents.

20 Preferably, according to the invention, the first step AA is carried out by one or more hypertext search engines using said first search criterium.

Preferably, according to the invention, in the second step BB, documents obtained from step A are semantically analysed up to a pre-set hypertextual level.

25 Advantageously, according to the invention, said first search criterium provides the use of keywords relevant to the content and/or the title of the documents, and/or the use of the definition of a context, and/or the use of the number of the following surfing levels and/or the use of the identification of the search engines to be used.

30 Preferably, according to the invention, second search criterium at least partially uses the keys and the context of the first search criterium.

Advantageously, according to the invention, search uses information relevant to the keywords and to the results of previous searches.

35 According to the invention, the method can further comprise the step of:

CC. displaying the documents obtained in the first search step on a graphic user interface, said graphic interface comprising a first

displaying window with the documents placed listed and a second window for drafting new documents.

Preferably, according to the invention, said one or more peripheral computers are connected to a server.

5 Preferably, according to the invention, one or more of said databases are stored within said server, said databases being updated on the peripheral computers.

Advantageously, according to the invention, first search criterium comprises GRID options.

10 According to the invention, method can further comprise the following step:

DD. communication of the first search criterium from the peripheral computer to the server.

15 According to the invention, the method can further comprise the step of:

EE. communication of the second search criterium from the peripheral computer to the server.

According to the invention, the method can further comprise the step of:

20 EF. carrying out a first search by the peripheral computer, by consultation of a local catalogue of the search contexts and of the already performed searches, as well as of the document addresses.

According to the invention, the method can further comprise the step of:

25 GG. retrieving of the documents obtained in the first search by the server.

According to the invention, the method can further comprise the step of:

30 HH. providing credential and keys, as well as initialization of communication channels by the server, suitable to enable the peripheral computer to p2p communications with other peripheral computers for retrieving and exchanging said documents.

Advantageously, according to the invention, p2p communications uses semi-private key cryptography.

35 According to the invention, a markup can be added to the retrieved documents, preferably a HTML markup.

According to the invention, the method can further comprise the step of:

II. carrying out an analysis of the documents according to said first search criterium by the peripheral computer.

5 According to the invention, the method can further comprise the step of:

JJ. carrying out an analysis according to said second search criterium of the documents discarded according to said first search criterium by the peripheral computer.

10 According to the invention, the method can further comprise the step of:

KK. carrying out an analysis according to said second search criterium of the documents discarded according to said first search criterium by the server.

15 According to the invention, the method can further comprise the step of:

LL. carrying out a search of new documents in the Internet by the peripheral computer or by the server.

20 According to the invention, the method can further comprise the step of:

MM. analysis, by the peripheral computer, according to said first search criterium of the documents obtained during the surfing.

According to the invention, the method can further comprise the step of:

25 NN. sending to the server the documents discarded during the analysis.

According to the invention, the method can further comprise the steps of:

30 NNN'. analysis by the server of the documents discarded according to said second search criterium,

OO. analysis by the server of the rejected documents according to said second search criterium.

According to the invention, the method can further comprise the step of:

35 PP. displaying, by the peripheral computer, through the user interface, the documents obtained from the search.

Preferably, according to the invention, said documents obtained from the search are **editable** on said peripheral computer.

Advantageously, according to the invention, the document(s) selected through the user interface are displayed on a window, and at the 5 same time a window is displayed to modify the local documents and the access to local databases.

Preferably, according to the invention, final documents are drafted in the XML format.

According to the invention, the method can further comprise the 10 step of:

QQ. adding by the server and/or peripheral computer to one or more of said data bases, one or more new documents created on the basis of all or part of the documents obtained from the search.

Preferably, according to the invention, a document hypertext 15 drafting mode can be selected by a user interface on the peripheral computer.

Advantageously, according to the invention, the list of the documents available for consultation and use is displayed.

According to the invention, one or more documents created on 20 the basis of all or part of the documents obtained from the search can be published on the Internet.

Preferably, according to the invention a hypertext surfing mode can be selected by a user interface on the peripheral computer.

Preferably, according to the invention the hypertext surfing 25 mode comprises the enabled of the automatic detection of the presence of the search keys and of the use of libraries.

Preferably, according to the invention a document search mode can be selected by a user interface on the peripheral computer.

Preferably, according to the invention, OLE-CLI libraries with 30 reader function on all the not HTML and not XML documents are used.

It is further **object-subject-matter** of the present invention a user or client peripheral computer, characterised in that it carries out step **AA** of the method according to the invention.

According to the invention, peripheral computer can carry out 35 step B of the method according to the invention.

It is further specific subject-matter object of the present invention a server, characterised in that it carries out step GG of the method according to the invention.

5 It is further object subject-matter of the present invention a computer program characterised in that it comprises code means suitable to carry out, when operating on a computer, step AA of the search, drafting and hypertext editing method according to the invention.

10 According to the invention, computer program can comprise code means suitable to carry out, when operating on a computer, step BB of the search, drafting and hypertext editing method according to the invention.

It is further object of the present invention a memory support readable by a computer, having a program stored, characterised in that the program is the computer program according to the invention.

15 It is further object subject-matter of the present invention a computer program characterised in that it comprises code means suitable to carry out, when operating on a computer, step GG of the search, drafting and hypertext editing method according to the invention.

20 It is further specific object subject-matter of the present invention a memory support readable by a computer, having a program stored on it, characterised in that the program is the computer program according to the invention.

25 The present invention will be now described, for illustrative but not limitative purposes, according to its preferred embodiments, with particular reference to the figures of the enclosed drawings, wherein:

figure 1 shows the block diagram of a first preferred embodiment of the method according to the invention;

figure 2 shows the block diagram of a second preferred embodiment of the method according to the invention;

30 figure 3a shows the first portion of the flow chart of the second preferred embodiment of the method according to the invention;

figure 3b shows the second portion of the flow chart of the second preferred embodiment of the method according to the invention;

35 figure 4a shows the first portion of the flow chart of a third preferred embodiment of the method according to the invention;

figure 4b shows the second portion of the flow chart of the third preferred embodiment of the method according to the invention.

The method according to the invention will be in the following indicated as "meta-engine" since, among the others features, it carries out searches among the results obtained from Internet search engines such as Google and Yahoo!.

5 According to a preferred embodiment of the invention, as illustrated in figure 1, one or more peripheral computers 300 are provided, to which the user can have access. On each one of the peripheral computers one database 200 is provided, said database being updated by the provider 100 (see continuous arrows).

10 It must be evident that the method according to the invention operates also with a single peripheral computer, since the method does not require a particular number of computers, being it a general-purpose method. In this case, i.e. when a single peripheral computer is present, the method according to the invention only permits advanced searches, 15 hypertext drafting and editing starting from the local databases.

Database refers to a collection of data (DC) by pointers. In this way data collection can be very compact, not being necessary that data are **stored** in different "pages", but one after the other.

20 For example, a general catalogue can be present in the database with the titles of the documents, said titles pointing to a specific portion of the DC content.

25 Data base can advantageously further comprise:
a list the treated contexts,
contextualised dictionaries, i.e. specialised dictionaries relevant to specific contexts;
semantic reference assemblies relevant to the contexts,
catalogue of the searches already carried out,
list and originary data content of the already consulted pages,
hypertext drafting control sequence starting from the raw text,
30 user registry,
enabled user list,
list of enabled contexts and the context / user matching.

Updating of database and DC can occur by the periodic providing of a **storage** support containing the same data.

35 Thus the user can carry out searches and editing documents, either texts, music, videos or other data assemblies, comprising textual complementary information suitable for their identification.

User inputs search keys, and the search engine associated with the database carries out the search, examining the results by a semantic analyser to decide which are relevant to the input keys.

Finally, the user has the possibility of accepting or discarding 5 those results of the second step not deemed to be relevant.

Search results are stored within the peripheral computer 10 memory, said computer analysing them to provide information to the provider (dashed arrows) about the required and not found documents, so as the provider can update the data base inputting one or more already asked but not found documents.

In a second embodiment of the invention, database can also be in a "server", to which the peripheral computers are connected.

In this case, meta-engine includes, at its basis, a peripheral 15 module installed on the "Client" computer of each meta-engine user, along with:

a module for managing and cataloguing the stored documents and local users master, i.e. the master of the users accessing the peripheral computer;

20 a module for managing communications, on a communication protocol, relevant to the received, processed and/or sent material, and relevant to the processing methods,

a module for managing the peer-to-peer authenticated communications (p2p) between users and for counting quotations and crossed uses and the relevant intellectual property rights.

25 Searches are carried out delegating to the known search engines (such as Google and Yahoo!), to refine then the search among found files surfing up to set hypertextual levels.

Basic function of the meta-engine according to the first 30 embodiment is that of "advanced textual editor" with communication from client to server (cts) and p2p, and output in HyperText Mark-up Language" (HTML) and eXtensible Markup Language (XML).

Selection of material and surfing of the hypertext links, as well 35 as format conversion can be entrusted to a server, but it will be more convenient entrusted to the Client, as it will be described in the following.

Obtained pages are resident on Client, Server maintaining instead originals and assemblies of operations to create final documents.

Implementation of said functions occurs by some user graphic interfaces and by modules, that, with respect to the ISO-OSI architecture, cover presentation layers and application layers.

5 Communication with other modules of the Client package, as well as with the server and with p2p makes reference to a communication protocol.

10 Preferably, said communication protocol is a not standard protocol, according to the invention being really advantageous using a protocol different from the communication protocols available on the market.

15 Basic technological ideas for this approach are:
maintaining the document history,
centralisation of different functions and reduction of the size of the additional software installed outside the meta-engine, delegating the more standard functions to the existing packages,

20 use of p2p technology as well as of the Client-Server technology, separation of the formal textual information or "markup" (complementary information useful, for instance, to create a hypertextual document starting from the ASCII text) from the informative basic content (ASCII text) and evaluation of resulting documents,

25 use of proprietary communication protocols,
acquisition of documents locally created on the data base using techniques different from HTML/XML by the local use of readers or of the "Object Linking and Embedding" (OLE) Method, specialised, separated and locally present, and/or by the sending to the Server.

30 Making reference to figures 3a and 3b, meta-engine during step 1 presents a first user interface for acquisition of searching criterium (prescinding from the other functionalities), followed by functionality of searching in the local copy of the catalogue of documents relevant to the already completed searches and of the searches under execution.

35 Thus, catalogue, in this specific case a network catalogue, i.e. a document shared by all the Clients to which the users can access, and is updated from the Server according to modes that can be set in the specific applications, e.g. by a Client control and relevant to the sole context of last set search.

Client carries out a first search on said catalogue, as well as on all the documents to which it can have local access.

Search in step 2 is a specific keywords search, to which the use of context keywords can follow. This kind of search is here referred as **1st level search**.

5 In case something relevant to the searched matter emerges from the search carried out, it is presented to the user through a second interface and the user selects the one it prefers as basis asking it to the server (if it is not already present in locale), communicating with one or more different Clients, certifying the relevant authorisations to the communication and providing to all the involved Clients the keys for 10 decrypting information, i.e. accredits for possible p2p communications.

Meanwhile, server carries out in a step 3 both a specific analysis on keys and an analysis of the context, thus starting a second level analysis **by repeating** of the first level search carried out by the Client.

15 Starting from the list of all the already carried out searches in the specific context, it starts a search in the central catalogue and in the cache of the pages obtained from the previous searches.

20 This allows providing to the Client all the raw pages previously accepted for the chosen context search; these pages are filtered again on the required keys (for example keywords and context).

Second level analysis is carried out, for example, in case keys are words, looking at the frequency of these words and comparing said frequency with the frequencies characteristics for the chosen context.

25 Now, in step 4, Client checks the index, which is common for all the Clients, of searches relevant to all the contexts for which he is enabled, said index containing also Client address where physically are the information (documents) to be found.

30 Thus Client asks to the Server, during a step 5, to find said documents, if the user interrupted the working session with the server, or if the Client where the information must be found is not connected to the network. This is done to optimise the times.

Otherwise, server accredits the Client for p2p communication as specified in the above.

35 Once terminated these documents finding operations, during a step 6, Client receives all the documents found during the search, obviously those available, and naturally has available local information not yet shared on the network.

Now Client can carry out during a step 7 a first semantic analysis of the above-mentioned documents and information, using the above specific and context keywords, and this can be done for all the contexts in such a way to make an "exhaustive" search.

5 The documents that were not judged pertinent at the end of said first semantic analysis are transmitted to the server for a step 8 wherein the latter effectuates a finer second level semantic analysis.

10 Said second semantic analysis includes the application of matrix, stylometric and statistic methods, such as the use of all the contexts known to the server, control in each available document of words being part of the same context, and the use of stylometric classes (occurrence frequencies and length of the words).

15 Documents that are thus deemed pertinent are transmitted to the Client, that, during a step 12, adds them to those already chosen and subjects the assembly of said documents to the user's selection.

At the same time of the second level analysis of step 8, another operation can be carried out during a step 9, in this case by the Client.

20 Said step 9 operation is a new search in the Internet of new documents according to the input search criterium. Search is delegate to the standard Internet search engines.

Results of this further search are transferred to a first semantic analysis, carried out by the Client during step 10, and then to a 2nd level semantic analysis, carried out by the Server during step 11.

25 During said 2nd level analysis, pages obtained from said searches are navigated by the Server at least up to link 10 level and are semantically analysed on text and context keys to establish if pages are relevant or not, to be then suggested to the user as blocks, preferably of about 1 megabyte or 20 URL, containing pages at the first link level and connected by links up to the maximum permitted level or up to the level 30 where no more links are present.

Finally, thus selected documents are suggested to the user during step 12 for final selection. He/she can accept or can refuse pages both at the abstract level, created by the same meta-engine employing specific algorithms, or at the full text level.

35 Abstract is created by a markup analysis and by comparison with documents of the same context. It is also possible to make use of frequency criterium of the words according to the context or, more roughly,

to the extraction of the first n words (usually between 20 and 50) from the textual context of the document found.

5 It is clear that other criterium can be added to the above ones, such as the presence of context keywords or the average length of the words, as well as the bibliographic references. Abstract creation procedure is indicated as "abstracting".

Now, the user can notify that he/she stops the search. In case of interruption, server sends to the Client all the documents already found.

10 Finally, user can intervene on the final document, making his/her modifications.

Accepted pages are transferred to the true editor, that during step 15 handles the content on order modifications, i.e. standard drafting operations allowing at the end to add a HTML markup or another program supporting the pure text scripting, thus creating control sequences relevant to the carried drafting operations and data allowing to regenerate the result starting from the original data.

20 Obviously, it is possible to save pages directly on the hard disk and then editing the same outside the advanced editor program by any program chosen by the user and sharing or not the final result.

25 In case of publication, Client sends the final edited XML document to the Server, said server publishing it in step 16.

It is to be added that, from peer-to-peer, beside the required documents, directly managed by the Client module, but not necessarily by the user presence, through which the users can exchange both other not catalogued documents, and observations and discussions, that can be shared and/or interconnected with other documents on demand.

30 In a third embodiment of the meta-engine, it further comprises: full functionality of hypertextual surfer or "browser"; functionality of removal of HTML instructions or "HTML stripping";

35 cryptography, preferably with a semiprivate key; introduction of "GRID" functions in the communication protocol, finalised to the transmission of semifinished data assemblies by the different peripheral computers toward the Server (to reduce the server work, each client carries out before hand a part of the second level analysis);

all OLE-CLI and OOP libraries relevant to the main applications (such as, for example, Word, Excel, Adobe, Macromedia, Flash and Director) link-edited in a static way or as Dynamic Link Libraries) (DDL);

5 possibility of memorising the whole in a XML database or on the server or on the Client to be able to re-use the information and parcel out the counting of the citations and of the uses, and to have direct access to the OLE libraries of Adobe PDF for print drafting and editing.

10 Each one of these performances is realised by a module locally operating on the Client and, when it is necessary, calls the "listener" asp/php of the server to transmit or to receive blocks of request or of data.

Within the Client package it is also present a compression/expansion LZ module.

Making reference to figures 4a and 4b, meta-engine has four use modes:

15 editor,
search,
extended "surfing", where link chains are automatically surfed,
simplified surfing, wherein link chains are not automatically surfed, but time by time after a control from the user.

20 Thus, meta-engine full use procedure (that will have to support the multi-allocation on independent windows beside tabbing) is as follows:

25 during a step 20, from the first interface, user chooses if using it according to a search mode activating step 23, editor mode activating step 21, or surfing, activating step 22; surfing could be extended (multiclass documents – i.e. documents recalling various modules handling different classes – markup, stripping and tabbing) or simplified (that will be the sole available mode if the program is called replying to a HTML, XML or Java link, requested by another program);

30 in case user has chosen:
editor mode, in step 21 the user sets a local path, or the title, or the number of a search; in step 22 a document list is presented to the user, the not complete documents first and the last one edited first; after his/her choice, during step 25, required document is delivered open in a window on the left side and a hypertext connection button or "tab" for each block not completely used as yet for listing the blocks, while during the contemporaneous step 26 a window on the right is opened, said right window having also the possibility of reading local documents drafted in

any format and to catalogue within the data base (DB) and include them in the data collection (DC);

5 Simplified surfing mode, in step 22a the user sets a **web** address, component enables HTML links in such a way that pages can be open in the same window or in a different window, using the OLE-CLI libraries only as reader function for all the not HTML and not XML documents (in case of Unix, since the syntactic analyser or "parser" XML cannot be integrated, XML will be considered as external format);

10 Extended surfing mode, during step 22b the user sets a web address; automatic **detection** of presence of keywords is enabled and OLE-CLI libraries are used also for markup, stripping and abstracting;

15 Search mode, during step 27, the user sets the search keywords; a second interface is suggested, wherein the context, the possible title, the number of depth levels, GRID options and engines to be interrogated, are chosen.

In case of the search, further important steps are present:

20 in a step 28, it is interrogated a local search catalogue and document context indicated on the basis of keywords and of the possible title and a third interface can be proposed with a list of titles during a step 30; at the same time

in a step 29 the assembly of parameters is communicated to the server through packages updating the **local** catalogue and the possible results are added to the user interface,

25 if in step 30 documents already included in the catalogue are chosen,

 - in step 31 the server

 - gets from the network the found documents,

30 - authenticates and accredits **p2p communications** for receiving original documents and the different sections, taking note of those not available,

 - extracts documents from data base sending to the Client control sequence for composition of documents,

 - extracts from the cache all the pages not available on the network, which were part of the previous final document,

35 - updates the cache copies of all that is still public and available by a comparison method, so that in case of excessive modifications, a

new "Uniform Resource Locator" URL is generated, sending then the relevant updating controls to the catalogue,

- sends to the Client all the parts resulting not publicly available or ~~too much extensively~~ modified of the identified documents in the

5 compressed and encrypted document of its library (cache copy),

- according to set GRID parameters, starts or not the searching of new material in a further step 32 on the engines or prepares a transparent process to the system supervisor/ "thread" listener and semantic analyst for what returns to the Client,

10 on its side, client meta-engine

- during step 33, activates different p2p connections discharging material, opening a tab for each block, decompressing and decrypting the whole at the moment of displaying,

- if it has chosen to locally carry out GRID advanced functions,

15 during a step 34,

- activates searches on engines,

- autonomously surfs the link chains,

- analysis at the first level pages obtained from the previous chains and searches, and

20 - for those accepted, communicates the summary markup and processed content to the server,

- those rejected are wholly sent natively to the server for further analysis levels,

25 - handles the editorial editing operations carried out on the documents by the user.

When he does not want any more material, the user pushes a suitable button on the interface thus interrupting the searches (both its own and the one of the server); however, the server in any case sends all those pages which both are already finished and have passed the more inner semantic analysis levels, updates the data bases and de-activates a specific thread listener, thus interrupting the searches.

30 For completion of the final document, the user inserts the markup closing the thread Clients, Client notifies the server and then puts the document, or does not put the document, for free catalogue or for payment consultation. Thread Clients toward inside (as for example those relevant to the drafting) depend on the presence of the user, those outside (as for example those relevant to the finding in the Internet of the

documents) not, all being resumable at the last registered point in case the Client is switched off.

5 Pages passing the semantic examination at the first Client level (from now on reference is made to the flow chart of the second realisation step, starting from step 10, diagram not repeated in figures 4a and 4b) and are discarded from the user, are sent natively to the server to go to further automatic analysis and eventually not automatic levels.

10 In case of acceptance at the subsequent analysis levels (for example determined by the belonging to the acceptable stylometric classes or by the belonging to link chains wherein accepted pages or in any case potentially acceptable pages are present), server proposes again them to the user in the following blocks, otherwise passes them to a not automatic analysis.

15 Once terminated the drafting of the new search document, the user can decide whether publishing it or not, and further can, for certain arguments, require a monitoring of a quarterly, semestral or annual duration of the new documents, as well as a count of the citations declared and of those implicit, and eventually of copies.

20 The present invention has been described for illustrative but not limitative purposes, according to its preferred embodiments, but it is to be understood that modifications and/or changes can be introduced by those skilled in the art without departing from the relevant scope as defined in the enclosed claims.

CLAIMS

1. Method of searching, drafting and editing of electronic files comprising the use of one or more peripheral computers or clients and a central computer or server, each client handling an assembly of one or 5 more databases, which refer to one or more datacollections by pointers and are updatable by the server and comprise electronic documents, said electronic documents comprising information suitable to identify the same documents, said databases assembly comprising one or more catalogues relevant to the documents, **the method being characterised in that** the 10 final documents obtained by the search or drafted on its basis are resident on client, the server maintaining instead ~~raw originay data content~~text, mark-up and commands for re-composing such final documents, the databases assembly comprising a history catalogue of the searches already carried out for each context by any client and an index of physical 15 locations of documents, the index being updated by the server, the method executing the following steps:

A. a Search step, in which:

20 A.1 the client performs a first-level search in the local copy of said history catalogue for the relevant context, and optionally
 A.2 the server performs a second-level search in said history catalogue and in the cache of pages obtained in previous searches by any client;

B. a documents Retrieval step, in which:

25 B.1 the client searches the location of the documents, found in step A, in said index of physical locations of documents,
 B.2 the client asks the documents to the clients to which they belong, by a p2p communication accredited by the server, or
30 B.3 the client asks the server to re-generate the documents, using ~~raw text~~originay data content, mark-up and commands;

C. a step of Semantic Analysis of the Results, in which

35 C.1 the client performs a first-level semantic analysis on the results obtained in step B, and optionally
 C.2 the server performs a second-level semantic analysis on the results rejected by the analysis of the client;

D. a step of Storing and Updating, in which the server stores the results of the search, included the documents drafted later on their basis, in the form of composition commands and ~~new-originary data content~~ ~~text~~ and mark-up, the server periodically updating all said catalogues, databases assembly and datacollections and said index of physical location of documents using the information relevant to the performed search.

5 2. Method according to claim 1, characterised in that said one or more catalogues comprise a list of title of the documents.

10 3. Method according to claim 1 or 2, characterised in that said one or more catalogues comprise a list of the contexts for which the documents are available, including the titles of the contexts.

15 4. Method according to one of preceding claims 1 — 3, characterised in that a first search criterium is used in step A.1 and a second search criterium is used in step A.2, both criteria using keywords and contexts.

5. Method according to claim 4, characterised in that said second search criterium is established by the server taking into account said first search criterium.

20 6. Method according to claim 4 or 5, the semantic analysis of step C.1 utilises the search criterium of step A.1.

7. Method according to one of preceding claim 4 — 6, characterised in that, in step C, it employs specialised dictionaries relevant to specific contexts ~~are employed in step C, relevant to specific contexts~~, and/or reference semantic assemblies relevant to the contexts.

25 8. Method according to one of the preceding claims 1 - 7, characterised in that said assembly of one or more databases is identical for all the clients.

9. Method according to one of the preceding claims 1 - 8, characterised in that said information suitable to identify documents are text information.

30 10. Method according to one of the preceding claims 1 - 9, characterised in that said documents are hypertext documents.

35 11. Method according to claim 10, characterised in that step A is carried out by one or more hypertext search engines.

12. Method according to claim 10 or 11, characterised in that, in step C, documents obtained from step A are semantically analysed up to a pre-set hypertextual level.

13. Method according to one of the preceding claims 4 – 12, 5 characterised in that said first search criterium provides the use of keywords relevant to the content and/or the title of the documents, and/or the use of the definition of a context, and/or the use of the number of the following surfing levels and/or the use of the identification of the search engines to be used.

10 14. Method according to one of the preceding claims 1 - 12, characterised in that semantic analysis of step C comprises an "abstracting" step.

15 15. Method according to one of the preceding claims 11 - 14, characterised in that documents are analysed in step C at least up to the third hypertextual level.

16. Method according to claim 15, characterised in that documents are analysed at least up to the fifth hypertextual level.

17. Method according to one of the preceding claims 11 - 16, 20 characterised in that the method further comprises the step A.3, in which the client displays the documents obtained in step A on a graphic user interface, said graphic interface comprising a first displaying window with the documents placed listed and a second window for drafting new documents.

18. Method according to one of the preceding claims 1 - 17, 25 characterised in that first search criterium comprises GRID options.

19. Method according to claim 1 - 18, characterised in that p2p communications use semi-private key cryptography.

20. Method according to claim 19, characterised in that a markup is added to the retrieved documents.

30 21. Method according to claim 20, characterised in that the markup is a HTML or XML markup.

22. Method according to one of the preceding claims 1 - 21, characterised in that it further comprises a step C.3, subsequent to step C.1, in which the client carries out a search of new documents in the 35 Internet.

23. Method according to claim 22, characterised in that it further comprises a step C.4, subsequent to step C.3, in which the client

analyseis, according to said first search criterium, the documents obtained during the surfing.

24. Method according to claim 23, characterised in that it further 5 comprises a step C.5, subsequent to step C.1, in which the client sends to the server the documents rejected during the analysis of C.1, the server analysing in step C.2 such rejected documents.

25. Method according to one of the preceding claims 1 - 24, 10 characterised in that it further comprises a step E in which the documents obtained from the search are displayed by the client through a user interface.

26. Method according to claim 25, characterised in that said documents obtained from the search are editable on said client.

27. Method according to claim 26, characterised in that the 15 document(s) selected through the user interface are displayed on a window, and at the same time a window is displayed to modify the local documents and the access to local data bases.

28. Method according to claim 26 or 27, characterised in that final documents are drafted in the XML format.

29. Method according to one of the preceding claims 1 - 28, 20 characterised in that one or more documents created on the basis of all or part of the documents obtained from the search can be published on the Internet.

30. Method according to one of the preceding claims 10 - 29, 25 characterised in that OLE-CLI libraries with reader function on all the not HTML and not XML documents are used.

31. User or client peripheral computer, characterised in that it carries out step A.1 and/or B and/or C.1 of the method according to one of the claims 1 - 30.

32. Server computer, characterised in that it carries out step A.2 and/or C.2 and/or D of the method according to one of the claims 1 - 30.

33. Computer program characterised in that it comprises code 35 means suitable to carry out, when operating on a computer, step A.1 and/or B and/or C.1 of the search, drafting and hypertext editing method according to one of the claims 1 - 30.

34. Memory support medium readable by a computer, having a program **stored** on it, characterised in that the program is the computer program according to claim 33.

5 35. Computer program characterised in that it comprises code means suitable to carry out, when operating on a computer, step A.2 and/or C.2 and/or D of the search, drafting and hypertext editing method according to according to one of the claims 1 - 30.

10 36. Memory support readable by a computer, having a program **stored** on it, characterised in that the program is the computer program according to claim 35.

ADVANCED METHOD OF SEARCHING, DRAFTING AND
EDITING OF ELECTRONIC FILES

ABSTRACT

5

The invention relates to a method of searching, drafting and editing of electronic files comprising the use of one or more peripheral computers, each computer handling an assembly of one or more updatable data bases comprising electronic documents, said electronic documents comprising information suitable to identify the same documents, each peripheral computer handling searches for, drafting of, communication about, and editing of, documents, the method being characterised in that said data base assembly provides one or more catalogues relevant to the documents, the search of documents or of part of documents being carried out employing at least one search criterium comprising the use of one or more catalogues.

The invention further relates to apparatuses and instruments necessary for carrying out the method according to the invention.

20